动 物 学 研 究 2002, Aug. 23 (4): 273 ~ 279 CN 53 - 1040/Q ISSN 0254 - 5853
Zoological Research

四个烙铁头蛇毒凝集素样蛋白基因的克隆与序列分析

魏 勤[1,2], 李 瑞[1], 王婉瑜[1], 熊郁良[1,3]
(1. 中国科学院昆明动物研究所, 云南 昆明 650223; 2. 中国科学院研究生院, 北京 100080)

摘要: 从烙铁头蛇(*Trimeresurus mucrosquamatus*)的毒腺中提取 mRNA, 利用 RT - PCR 进行体外扩增, 获得凝集素样蛋白基因, 克隆至 PMD18 - T 载体中, 筛选出 4 种凝集素样蛋白基因(命名为 TML-1、TML-2、TML-3 和 TML-4)。由基因序列推导出的氨基酸序列表明: TML-1, 2, 3, 4 序列中均有 CRD 结构。序列同源性比较和 Cys 位点分析推测: TML-1 和 TML-2 可能分别是类似于 flavocetin-A 的蛇毒凝集素样蛋白的 α 亚基和 β 亚基; TML-3 可能类似于 GPIb-bp 的蛇毒凝集素样蛋白的 α 亚基, TML-4 则可能是类似于Ⅸ/Ⅹ-bp 的蛇毒凝集素样蛋白的 β 亚基。

关键词: 烙铁头; 凝集素样蛋白; cDNA; 序列分析
中图分类号: Q959.62; Q75 **文献标识码**: A **文章编号**: 0254 - 5853(2002)04 - 0273 - 07

Cloning and Sequence Analysis of cDNAs Encoding Four C-type Lectin-like Proteins from Snake *Trimeresurus mucrosquamatus*

WEI Qin[1,2], LI Rui[1], WANG Wan-yu[1], XIONG Yu-liang[1]
(1. *Kunming Institute of Zoology, the Chinese Academy of Sciences, Kunming* 650223, *China*;
2. *Graduate School of the Chinese Academy of Sciences, Beijing* 100080, *China*)

Abstract: The total mRNA was prepared and purified from the venom glands of snake *Trimeresurus mucrosquamatus*. The cDNAs encoding C-type lectin-like proteins were amplified by RT - PCR and cloned into the PMD 18 - T vectors. The positive clones encoding four C-type lectin-like proteins (TML-1, TML-2, TML-3 and TML-4) were selected and sequenced. Amino acid sequences of these proteins were deduced and each contains carbohydrate recognition domain (CRD). Homology comparison and cysteine position analysis indicate that TML-1 and TML-2 might be the α and β subunits of a C-type lectin-like protein similar to flavocetin-A, TML-3 the α subunit of a C-type lectin-like protein similar to GPIb-bp, and TML-4 the β subunit of a C-type lectin-like protein similar to Ⅸ/Ⅹ-bp.

Key words: *Trimeresurus mucrosquamatus*; C-type lectin-like protein; cDNA; Sequence analysis

蛇毒中含有许多作用于血液系统的蛋白质。其中一类蛋白质的氨基酸序列中含有与动物凝集素的糖识别域(carbohydrate recognition domain, CRD)相似的特征序列。与蛇毒凝集素不同,这类蛋白质一般不能与糖结合,因而被称为蛇毒凝集素样蛋白(venom C-type lectin-like protein)。蛇毒凝集素样蛋白一般是由 2 个相似亚基组成的异构二聚体,分子量约为 30 kDa。但也有少数蛇毒凝集素样蛋白是由多个异构二聚体组成的多聚体,分子量在 50 ~ 150 kDa,如 flavocetin-A (Taniuchi *et al*., 1995)和 convulxin (Francischetti *et al*., 1997)。从蛇毒中分离到的凝集素样蛋白的氨基酸序列同源性大致在 30% ~ 90%。虽然其结构非常相似,但生物活性和作用机制各不相同。一些蛇毒凝集素样蛋白通过结合凝血因子Ⅸ或Ⅹ起到抗凝作用,如从黄绿烙铁头(*Trimeresurus flavoviridis*)蛇毒中分离得到的Ⅸ/Ⅹ-

收稿日期: 2002 - 02 - 09; 接受日期: 2002 - 03 - 15
3. 通讯作者 (Corresponding author), Fax: 86 - 871 - 5191823, Tel: 86 - 871 - 5196747

bp(Atoda *et al*.,1991)。而从美洲矛头蝮(*Bothrops jararaca*)蛇毒中得到的 *bothrajaracin* 则是一个凝血酶抑制剂(*Zingali et al*.,1993)。一些蛇毒凝集素样蛋白通过作用于 von Willebrand 因子或 GPIb-Ⅸ-Ⅴ、GPVI 等血小板受体而引起血小板聚集,如 botrocetin (Fujimura *et al*., 1991)、convulxin (Leduc *et al*., 1998)和 rhodocytin (Shin & Morita, 1998)。但 flavocetin-A (Taniuchi *et al*., 1995)和 echicetin (Peng *et al*., 1993)等蛋白却是通过与 GPIb-Ⅸ-Ⅴ的结合而阻断其与 von Willebrand 因子的结合而抑制血小板聚集。但是研究蛇毒凝集素样蛋白的结构依然是研究其功能的基础。

近年来,已从南美恐怖响尾蛇(*Crotalus durissus terrificus*)(Leduc *et al*., 1998)、黄绿烙铁头(*Trimeresurus flavoviridis*)(Shin *et al*.,2000)、红口蝮(*Calloselasma rhodostoma*)(Chung *et al*.,1999)、日本蝮(*Agkistrodon halys blomhoffii*)(Sakurai *et al*., 1998)、尖吻蝮(*Deninagkistrodon acutus*)(Chen *et al*.,2000)等蛇中克隆出一些蛇毒凝集素样蛋白的基因。但是,尚未见有从烙铁头蛇(*Trimeresurus mucrosquamatus*)中克隆凝集素样蛋白基因的报道。我们根据蛇毒凝集素样蛋白的基因 5′端和 3′端非翻译区的保守序列,设计简并引物,从烙铁头蛇毒腺中克隆到 4 个凝集素样蛋白基因,现将这一结果报道如下。

1 材料与方法

1.1 材 料

烙铁头蛇采自湖南沅陵。mRNA 提取试剂盒及 cDNA 合成试剂盒购自 Promega 公司, PCR 反应试剂盒、PMD18-T 载体、*E.coli* JM109 感受态宿主菌均购自 Takara 公司。

1.2 烙铁头毒腺 mRNA 的提取

切下蛇头,并立即将蛇头放入液氮中待用。剥离毒腺,按照试剂盒说明提取 mRNA。

1.3 cDNA 的合成

按照试剂盒说明进行 cDNA 的合成。

1.4 寡核苷酸引物的合成

根据不同来源蛇毒凝集素样蛋白基因保守的 5′端和 3′端非翻译区序列 (Leduc & Bon, 1998; Shin *et al*., 2000),设计出正向引物 P5: 5′-AGGGAAGGAAGGAAGACCATGG-3′; 反向引物 P3: 5′-GGGGGCTTCCTTGCTTCTCCAG-3′,交由 Takara 公司合成。

1.5 目的 DNA 的获得

以 cDNA 为模板,进行 PCR 体外扩增。反应条件: 94 ℃ 5 min; 然后 94 ℃变性 30 s, 52 ℃退火 40 s, 72 ℃延伸 1 min, 共 35 个循环; 72 ℃保温 10 min。

1.6 目的 DNA 的重组克隆

将 PCR 产物直接克隆到 PMD18-T 载体上,转化 *E.coli* JM109 感受态细胞。37 ℃下,置于含氨苄青霉素 (100 μg/mL)、IPTG 和 X-gal 的 LB 平板上培养。挑取白色菌落进行 PCR 鉴定。

1.7 DNA 序列测定及分析

PCR 阳性克隆由上海基康生物技术有限公司进行序列测定。采用 Vector NTI Suite 6.0 软件进行序列比较分析及分子量计算。

2 结果和讨论

2.1 cDNA 的 PCR 扩增和克隆

根据不同来源蛇毒凝集素样蛋白基因 5′端和 3′端非翻译区的保守序列设计引物。经 PCR 扩增得到 500 bp 左右的产物,克隆到 PMD18-T 载体上,转化大肠杆菌 JM109。然后经蓝白斑筛选出 25 个白色菌落。经 PCR 鉴定,其中 10 个克隆插入了目的 DNA 大小的片段。对这 10 个阳性克隆进行测序分析,获得 4 个不同的凝集素样蛋白基因,分别命名为 TML-1、TML-2、TML-3 和 TML-4。cDNA 序列及推导出的氨基酸序列见图 1。

2.2 序列分析

由 cDNA 序列推出的氨基酸序列表明,TML-1,2,3,4 均包含一段由 23 个氨基酸残基组成的疏水性极强的肽段,且具有很高的同源性(为 82.6%)。与其他蛇毒凝集素样蛋白的信号肽进行同源性比较,表明这一肽段为信号肽。TML-1 的成熟肽由 135 个氨基酸残基组成,分子量为 15.711 kDa。TML-2 的成熟肽由 123 个氨基酸残基组成,分子量为 14.309 kDa。TML-3 的成熟肽由 133 个氨基酸残基组成,分子量为 15.765 kDa。TML-4 的成熟肽由 121 个氨基酸残基组成,分子量为 14.283 kDa(图 1)。

TML-1,2,3,4 的成熟肽里均含有 CRD 结构,且与其他已报道的蛇毒凝集素样蛋白的氨基酸序列有很高的同源性(图 2,图 3)。一般来说,蛇毒凝集素样蛋白的 CRD 结构中 Cys 的位点是保守的。低分子量的凝集素样蛋白α亚基中,Cys位于4、15、32、

```
TML-1                                                               AGGGAAGGAAGGAAGACC
TML-2                                                               AGGGAAGGAAGGAAGACC
TML-3                                                               AGGGAAGGAAGGAAGACC
TML-4                                                               AGGGAAGGAAGGAAGACC

           1                                               50
TML-1      ATGGGGCGATTCACCTTCGTGAGCTTCGGCTTGCTGGTCGTGTTCCTCTCCCTGAGTGGTACTGGAGCTGATTTC
TML-2      ATGGGGCGATTCATCTTCGTGAGCTTCGGCTTGCTGGTCGTGTTCATCTCCCTGAGTGGAACTGAAGCTGGTTTC
TML-3      ATGGGGCGATTCATCTTCGTGAGCTTCGGCTTGCTGGTCGTGTTCCTCTCCCTGAGTGGTACTGGAGCTGATTGT
TML-4      ATGGGGCGATTCATCTTCGTGAGCTTCGGCTTGCTGGTCGTGTTCCTCTCCCTGAGTGGAACTGCAGCTGATTGT
            M  G  R  F  T  F  V  S  F  G  L  L  V  V  F  L  S  L  S  G  T  G  A  D  F
            M  G  R  F  I  F  V  S  F  G  L  L  V  V  F  I  S  L  S  G  T  E  A  G  F
            M  G  R  F  I  F  V  S  F  G  L  L  V  V  F  L  S  L  S  G  T  G  A  D  C
            M  G  R  F  I  F  V  S  F  G  L  L  V  V  F  L  S  L  S  G  T  A  A  D  C

                                  100                                               150
TML-1      GATTGTATCCCTGGTTGGTCCGCCTATGATCGGTATTGCTACCAGGCCTTCAGTGAACCGAAAAACTGGGAAGAT
TML-2      TGTTGTCCCTTGGGTTGGTCCTCCTATGATGAGCATTGCTACCAGGTCTTCCAACAAAAGATGAACTGGGAAGAT
TML-3      CCCTCTGATTGGTCCTCCTTTACACGGTATTGCTACAAGCCCTTTAAACAACTCAAGACCTGGGAAGATGCAGAG
TML-4      CCCTCTGATTGGTCCTCTTATGCAGGGCATTGCTACAAGCCCTTCAATGAACCAAAAAACTGGGCCGATGCAGAG
            D  C  I  P  G  W  S  A  Y  D  R  Y  C  Y  Q  A  F  S  E  P  K  N  W  E  D
            C  C  P  L  G  W  S  S  Y  D  E  H  C  Y  Q  V  F  Q  Q  K  M  N  W  E  D
            P  S  D  W  S  S  F  T  R  Y  C  Y  K  P  F  K  Q  L  K  T  W  E  D  A  E
            P  S  D  W  S  S  Y  A  G  H  C  Y  K  P  F  N  E  P  K  N  W  A  D  A  E
-
                                                           200
TML-1      GCAGAGAGTTTCTGCGAGGAGGGGGTGAAGACCTCGCATCTGGTCTCCATCGAAAGCTCCGGAGAAGGAGACTTC
TML-2      GCAGAGAAATTCTGCACACAACAGCACACAGGCAGCCATCTGGTCTCCTATGAGAGCAGTGAAGAAGTAGATTTT
TML-3      AGGTTCTGCTGGGAGCAGGTGAAAGGTGCGCATCTGGTCTCCATCGAAAGCTCCGGAGAAGGAGACTTCGTGGCC
TML-4      AATTTCTGCACACAACAGCACGCAGGCGGACATCTGGTCTCCTTCCAAAGCAGTGAAGAAGCAGATTTTGTGGTC
            A  E  S  F  C  E  E  G  V  K  T  S  H  L  V  S  I  E  S  S  G  E  G  D  F
            A  E  K  F  C  T  Q  Q  H  T  G  S  H  L  V  S  Y  E  S  S  E  E  V  D  F
            R  F  C  W  E  Q  V  K  G  A  H  L  V  S  I  E  S  S  G  E  G  D  F  V  A
            N  F  C  T  Q  Q  H  A  G  G  H  L  V  S  F  Q  S  S  E  E  A  D  F  V  V
-
                                  250                                               300
TML-1      GTGGCCCAGCTGGTCGCTGAGAAGATAAAGACGTCCTTTCAGTATGTCTGGATTGGGCTGAGGATTCAAAACAAA
TML-2      GTGGTCTCGAAGACCTTACCAATTTTGAAAGCCAGTTTTGTCTGGATCGGACTGAGCAATGTCTGGAATGCATGC
TML-3      CAGCTGCTCTCCGAGAACATAAAGACCACCAAATACCATGTCTGGATCGGACTGAGCGTTCAAAACAAAAGACAG
TML-4      AAGCTGGCCTTCCAAACTTTTGGCCACAGTATTTTCTGGATGGGACTGAGCAATGTCTGGAATCAATGCGACTGG
            V  A  Q  L  V  A  E  K  I  K  T  S  F  Q  Y  V  W  I  G  L  R  I  Q  N  K
            V  V  S  K  T  L  P  I  L  K  A  S  F  V  W  I  G  L  S  N  V  W  N  A  C
            Q  L  L  S  E  N  I  K  T  T  K  Y  H  V  W  I  G  L  S  V  Q  N  K  R  Q
            K  L  A  F  Q  T  F  G  H  S  I  F  W  M  G  L  S  N  V  W  N  Q  C  D  W
-
                                                           350
TML-1      GAACAGCAATGCAGGTCGGAGTGGAGCGATGCCTCCAGTGTCAATTATGAGAACTTGTTTAAACAATCTTCCAAA
TML-2      AGGTTGCAGTGGAGCGATGGCACCGAGCTTATGTACAATGCCTGGACTGCAGAATCTGAGTGTATCGCATCCAAG
TML-3      CAATGCAGGTCGATATGGAGCGATGGCTCCAGCGTCAGTTATGAGAACTTGGTTAAGCCATTTTCCAAAAAGTGC
TML-4      CAGTGGAGCAATGCTGCCATGCTCAGATACAAAGCCTGGGCTGAAGAATCTTACTGTGTCTACTTCAAGTCAACT
            E  Q  Q  C  R  S  E  W  S  D  A  S  S  V  N  Y  E  N  L  F  K  Q  S  S  K
            R  L  Q  W  S  D  G  T  E  L  M  Y  N  A  W  T  A  E  S  E  C  I  A  S  K
            Q  C  R  S  I  W  S  D  G  S  S  V  S  Y  E  N  L  V  K  P  F  S  K  K  C
            Q  W  S  N  A  A  M  L  R  Y  K  A  W  A  E  E  S  Y  C  V  Y  F  K  S  T

                                  400                                               450
TML-1      AAATGTTATGCGCTGAAAAAAGGGACAGAGCTTCGCACGTGGTTCAATGTTTACTGTGGAAGAGAAAATCCTTTC
TML-2      ACAACTGATAACCAATGGTGGAGTATGGACTGCAGCAGTAAACGCTATGTCGTCTGCAAGTTCTAGGCATAGTCT
TML-3      TTTGTGCTGAAAAAAGAGTCAGAGTTCCATAAGTGGTTTAATATTTACTGTGGAGAACGAAATCTTTTCATGTGC
TML-4      AATAACAAATGGAGGAGTAGAGCCTGCAGAATGGAGGCACATTTCGTCTGCGAGTTCCAGGCATAGTCTGAAGAT
            K  C  Y  A  L  K  K  G  T  E  L  R  T  W  F  N  V  Y  C  G  R  E  N  P  F
            T  T  D  N  Q  W  W  S  M  D  C  S  S  K  R  Y  V  V  C  K  F
            F  V  L  K  K  E  S  E  F  H  K  W  F  N  I  Y  C  G  E  R  N  L  F  M  C
            N  N  K  W  R  S  R  A  C  R  M  E  A  H  F  V  C  E  F  Q  A
-

TML-1      GTCTGCAAGTACACGCCAGAATGTTAAGATCCAGCTGAGTGAAGTCTGGAGAAGCAAGGAAGCCCCC
TML-2      GAAGATGCAGCTGTCTGAAGTCTGGAGAAGCAAGGAAGCCCCC
TML-3      AAGTTCCTGCAACCGCGTTAAGATCCAGCTGTGTGAAGTCTGGAGAAGCAAGGAAGCCCC
TML-4      GCAGCTGAGTGAAGTCTGGAGAAGCAAGGAAGCCCCC
            V  C  K  Y  T  P  E  C

            K  F  L  Q  P  R
```

图 1 烙铁头蛇毒凝集素样蛋白 TML-1、TML-2、TML-3 和 TML-4 的 cDNA 序列及推定的氨基酸序列

Fig.1 Complete nucleotide sequences and deduced amino acid sequences of TML-1, TML-2, TML-3, and TML-4

核酸序列方向为由 5′端到 3′端,核酸序列的编号(1)从起始密码子(ATG)起始。核酸序列下为相应的氨基酸序列。斜体字母表示信号肽序列。

The nucleotide residues of coding region are numbered in the 5′ to 3′ direction. The numbering started at the initial code (ATG). Beneath the nucleotide sequence is the deduced amino acid sequence. The putative signal sequence is shown in italics.

α subunit

```
                                          1          10          20          30          40
TML-1         MGRFTFVSFGLLVVFLSLSGTGADFDCIPGWSAYDRYCYQAFSEPKNWEDAESFCEEGVKTSHLV
flavocetin-A  MERLIFVSFGLLVVILSLSGTGADFDCIPGWSAYDRYCYQAFSKPKNWEDAESFCEEGVKTSHLV
convulxin     MGRFIFVSFGLLVLFLSLSGTGAGLHCPSDWYYYDQHCYRIFNEEMNWEDAEWFCTKQAKGAHLV

                   50          60          70          80          90          100
TML-1         SIESSGEGDFVAQLVAEKIKTSFQYVWIGLRIQNKEQQCRSEWSDASSVNYENLFKQSSKKCYAL
flavocetin-A  SIESSGEGDFVAQLVAEKIKTSFQYVWIGLRIQNKEQQCRSEWSDASSVNYENLVKQFSKKCYAL
convulxin     SIKSAKEADFVAWMVTQNIEESFSHVSIGLRVQNKEKQCSTKWSDGSSVSYDNLLDLYITKCSLL

                110         120         130         Identity
TML-1         KKGTELRTWFNVYCGRENPFVCKYTPEC            100%
flavocetin-A  KKGTELRTWFNVYCGTENPEVCKYTPEC            96.3%
convulxin     KKETGFRKWFVASCIGKIPFVCKFPPQC            50.4%
```

β subunit

```
                                          1          10          20          30          40
TML-2         MGRFIFVSFGLLVVFISLSGTEAGFCCPLGWSSYDEHCYQVFQQKMNWEDAEKFCTQQHTGSHLV
flavocetin-A  MGQFIFVSFGFLVVATSLSGTEAGFCCPLGWSSYDEHCYQVFQQKMNWEDAEKFCTQQHKGSHLV
convulxin     MGRFIFVSFGLLVVFLSLSGSEAGFCCPSHWSSYDRYCYKVFKQEMTWADAEKFCTQQHTGSHLV

                   50          60          70          80          90          100
TML-2         SYESSEEVDFVVSKTLPILKASFVWIGLSNVWNACRLQWSDGTELMYNAWTAESECIASKTTDNQ
flavocetin-A  SFHSSEEVDFVTSKTFPILKYDFVWIGLSNVWNECTKEWSDGTKLDYKAWSGGSDCIVSKTTDNQ
convulxin     SFHSTEEVDFVVKMTHQSLKSTFFWIGANNIWNKCNWQWSDGTKPEYKEWHEEFECLISRTFDNQ

             110         120         Identity
TML-2         WWSMDCSSKRYVVCKF--        100%
flavocetin-A  WLSMDCSSKYYVVCKFQA        81.6%
convulxin     WLSAPCSDTYSFVCKFEA        61.6%
```

图 2 烙铁头蛇毒凝集素样蛋白 TML-1、TML-2 与其他高分子量蛇毒凝集素样蛋白质序列比较

Fig.2 Comparison of amino acid sequences of TML-1 and TML-2 with those of other high molecular weight snake venom C-type lectin-like proteins

点表示蛇毒凝集素样蛋白氨基酸序列里保守的半胱氨酸残基;加框表示额外的半胱氨酸残基。每条序列后面列出了同源性比较。序列 flavocetin-A 引自 Shin *et al*.(2000);convulxin 引自 Leduc *et al*.(1998)。

The numbering is based on the sequence of each protein. Dots indicate the position of well conserved cysteine residues. The additional cysteine residues are boxed. The identity is at the end of each sequence. Sequences of flavocetin-A from Shin *et al*. (2000) and convulxin from Leduc *et al*. (1998).

81、104、121、129 位;β 亚基中,Cys 位于 4、15、32、77、98、113、121 位。同时,Ⅸ/Ⅹ-bp(Mizuno *et al*., 1997)和 botrocetin(Sen *et al*., 2001)的二硫键位置已由晶体衍射法得到了确认:其 α 和 β 亚基各有 6 个 Cys 形成 3 个链内二硫键。此外,α 和 β 亚基里还各有 1 个 Cys 配对形成的 1 个链间二硫键。这些二硫键的具体构成方式为:α 亚基中,Cys4 – 15、Cys32 – 129和Cys104 – 121;β亚基中,Cys4 – 15、Cys32 –

α subunit

```
                                        1        10        20        30        40
TML-3       MGRFIFVSFGLLVVFLSLS—GTGA---DCPSDWSSFTRYCYKPFKQLKTWEDAERFCWEQVKGAHL
AL-B        ---------------------------DCPSDWSSFKQYCYQIIKQLKTWEDAERFCMDQVKGAHL
mamushigin  MGRFIFVSFGLLVVFLSLSGAEDDS--DCPSDWSSNGRFCYKLFQQKMKWADAERFCTEQRTGAHL
botrocetin  ---------------------------DCPSGWSSYEGNCYKFFQQKMNWADAERFCSEQAKGGHL
IX/X-bp     MGRFIFMSFGLLVVAASLRG—TGA---DCLSGWSSYEGHCYKAFEKYKTWEDAERVCTEQAKGAHL
X-bp        ---------------------------DCSSGWSSYEGHCYKVFKQSKTWADAESFCTKQVNGGHL

                 50        60        70        80        90       100
TML-3       VSIESSG-EGDFVAQLLSENIKTTKYHVWIGLSVQNKRQQCRSIWSDGSSVSYENLVKPFSKKCFV
AL-B        VSIESYR-EAVFVAQQLSENVKTTKYDVWIGLSVVNKGQQCSSEWSDGSSVSYENLVKPLSKKCFV
mamushigin  VSIESNT-EAAFVNQMISENIKKTDY-VWIGLTVQNEEQQCKSRWSDRSSVSYENLVKPNSKKCFV
botrocetin  VSIKIYSKEKDFVGDLVTKNIQSSDLYAWIGLRVENKEKQCSSEWSDGSSVSYENVVERTVKKCFA
IX/X-bp     VSIESSG-EADFVAQLVTQNMKRLDFYIWIGLRVQGKVKQCNSEWSDGSSVSYENWIEAESKTCLG
X-bp        VSIESSG-EADFVGQLIAQKIKSAKIHVWIGLRAQNKEKQCSIEWSDGSSISYENWIEEESKKCLG

              110       120       130        Identity
TML-3       LKKESEFHKWFNIYCGERNLFMCKELQPR        100%
AL-B        LKKGTEFRKWFNVACEQKHLFMCKFL---         76.7%
mamushigin  LKEYEGSRKWFNVYCGQKYAFMCKFLRPR         66.4%
botrocetin  LEKDLGFVLWINLYCAQKNPFVCKSPPP-         54.5%
IX/X-bp     LEKETDFRKWVNIYCGQQNPFVCEA----         60.2%
X-bp        VHIETGFHKWENFYCEQQDPFVCEA----         57.1%
```

β subunit

```
                                      1        10        20        30        40
TML-4       MGRFIFVSFGLLVVFLSLSGTAA--DCPSDWSSYAGHCYKPFNEPKNWADAENFCTQQHAGGHLVS
AL-B        -------------------------DCPSDWSSYDLYCYRVFQEKKNWERAEKFCTQQHTDSHLVS
mamushigin  MGRFIFLSFGLLVVFVSLSGTGA--DCPSDWSSYEGHCYRVFQKEMTWEDAEKFCTQQRKESHLVS
botrocetin  -------------------------DCPPDWSSYEGHCYRFFKEWMHWDDAEEFCTEQQTGAHLVS
IX/X-bp     MGRFIFMSFGFLVVFLSLSGTAA--DCPSDWSSYEGHCYKPS-EPKNWADAENFCTQQHAGGHLVS
X-bp        -------------------------DCPSDWSSYEGHCYKPFNEPKNWADAENFCTQQHTGSHLVS

               50        60        70        80        90       100
TML-4       FQSSEEADFVVKLAFQTFGHSIFWMGLSNVWNQCDWQWSNAAMLRYK--AWAEESYCVYFKSTNNK
AL-B        FDSSEEVDFVASKTFPVLKHDLVWIGLGSVWNACKLQWSDGTELKYN--AWSAESECITSKSTDNE
mamushigin  FHSSEEVDFVVSMTWPILKYDFVWIGLNNIWNECMVEWYDGTRLSHN--AWITESECIAAKTTQNQ
botrocetin  FQSKEEADFVRSLTSEMLKGDVVWIGLSDVWNKCRFEWTDGMEFDYDDYYLIAEYECVASKPTNNK
IX/X-bp     FQSSEEADFVVKLAFQTFGHSIFWMGLSNVWNQCNWQWSNAAMLRYK--AWAEESYCVYFKSTNNK
X-bp        FQSTEEADFVVKLAFQTFDYGIFWMGLSKIWNQCNWQWSNAAMLKYT--DWAEESYCVYFKSTNNK

              110       120       Identity
TML-4       WRSRACRMEAHFVCEFQA         100%
AL-B        WLTRSCSRTYPFVCKFQA          55.3%
mamushigin  WLSRPCSRTYNVVCKFQE          49.6%
botrocetin  WWIIPCTRFKNFVCEFQA          51.2%
IX/X-bp     WRSRACRMMAQFVCEFQA          95.9%
X-bp        WRSITCRMIANFVCEFQA          86.2%
```

图 3 烙铁头蛇毒凝集素样蛋白 TML-3、TML-4 与其他低分子量蛇毒凝集素蛋白质序列比较

Fig.3 Comparison of amino acid sequences of TML-3 and TML-4 with those of other low molecular weight snake venom C-type lectin-like proteins

点表示蛇毒凝集素样蛋白氨基酸序列里保守的半胱氨酸残基；下划线表示Ⅸ/Ⅹ-bp 序列里的钙离子结合位点。序列中加“-”便于与图2中氨基酸序列的编号对齐。每条序列后面列出了同源性比较。序列 AL-B 引自 Usami *et al*.（1996）；mamushigin 引自 Sakurai *et al*.（1998）；botrocetin 引自 Usami *et al*.（1993）；Ⅸ/Ⅹ-bp 引自 Matsuzaki *et al*.（1996）；Ⅹ-bp 引自 Atoda *et al*.（1998）。

The numbering is based on the sequence of each protein. Dots indicate the position of well conserved cysteine residues. Gaps（-）are introduced to improve alignment with the sequence number in Fig.2. Underlines indicate Ca^{2+}-binding sites of Ⅸ/Ⅹ-bp. The identity is at the end of each sequence. Sequences of AL-B from Usami *et al*.（1996），mamushigin from Sakurai *et al*.（1998），botrocetin from Usami *et al*.（1993），Ⅸ/Ⅹ-bp from Matsuzaki *et al*.（1996），and Ⅹ-bp from Atoda *et al*.（1998）.

121 和 Cys98-113；链间二硫键由 α 亚基的 Cys81 和 β 亚基的 Cys77 组成。目前所发现的低分子量蛇毒凝集素样蛋白主要可分为两大类：① 血小板膜糖蛋白 GPIb 结合蛋白(GPIb-bp)，如 AL-B(Usami *et al*.，1996)和 mamushigin (Sakurai *et al*.，1998)；②凝血因子结合蛋白，如Ⅸ/Ⅹ-bp (Matsuzaki *et al*.，1996)和Ⅹ-bp (Atoda *et al*.，1998)。其中，凝血因子结合蛋白与Ⅸ因子和Ⅹ因子的结合都依赖于钙离子。Mizuno (1997) 通过对Ⅸ/Ⅹ-bp 进行晶体衍射研究确定了其钙离子的结合位点(图 3)，这些位点在Ⅸ-bp、Ⅹ-bp 中也是保守的(Atoda *et al*.，1998；Mizuno *et al*.，1999)。TML-3 的 Cys 位点与低分子量蛋白的 α 亚基的 Cys 位点完全一致，而 TML-4 的 Cys 位点与低分子量蛋白的 β 亚基的 Cys 位点也完全一致。氨基酸序列同源性比较表明，TML-3 与 AL-B 和 mamushigin 的 α 亚基同源性较高(分别为 76.7%和 66.4%)，而 TML-4 与Ⅸ/Ⅹ-bp 和Ⅹ-bp 的 β 亚基同源性较高(分别为 95.9%和 86.2%)。而且，在 TML-4 的氨基酸序列里也存在钙离子结合位点。由此推测，TML-3 可能是烙铁头蛇毒里某个类似于 GPIb-bp 的凝集素样蛋白的 α 亚基，TML-4 可能是烙铁头蛇毒里某个类似于Ⅸ/Ⅹ-bp 的凝集素样蛋白的 β 亚基。

与低分子量凝集素样蛋白相比，高分子量的凝集素样蛋白，如 flavocetin-A (Shin *et al*.，2000)和 convulxin (Leduc *et al*.，1998)的氨基酸序列中除了拥有上述位点的 Cys 外，在 α 亚基的 C 端还有 1 个 Cys135，而 β 亚基的 N 端还有 1 个 Cys3。在它们的序列中没有发现钙离子结合位点，实验结果也证实其生物活性不依赖于钙离子(Taniuchi *et al*.，1995；Francischetti *et al*.，1997)。对 flavocetin-A 进行晶体衍射分析表明，flavocetin-A 为(αβ)₄ 结构，α 和 β 亚基各有 6 个 Cys 形成 3 个链内二硫键，α 亚基中 Cys81 与 β 亚基中 Cys77 形成链间二硫键，将 α 亚基和 β 亚基连接为 αβ 异构二聚体，每个 αβ 异构二聚体里 α 亚基的 Cys135 又与相邻的 αβ 异构二聚体的 β 亚基的 Cys3 形成 1 个二硫键，最终形成“首尾”相连的四聚体(Fukuda *et al*.，2000)。氨基酸序列同源性比较表明：TML-1 的 C 端有一个额外的 Cys135，TML-2 的 N 端也有一个额外的 Cys3。TML-1 的氨基酸序列与 flavocetin-A 的 α 亚基序列的同源性高达 96.3%；TML-2 序列与 flavocetin-A 的 β 亚基序列的同源性高达 81.6%。因此，我们推测 TML-1 和 TML-2 可能是烙铁头蛇毒里某个高分子量的凝集素样蛋白的 α 亚基和 β 亚基，且这个蛋白质也是由几个 αβ 异构二聚体里“首尾”连接而成。事实上这种推论已从我们的另一项工作结果(从烙铁头蛇毒中分离到这一高分子量的凝集素样蛋白)得以证实。在 TML-1 和 TML-2 的氨基酸序列中，与Ⅸ/Ⅹ-bp 的钙离子结合位点相对应的氨基酸残基发生了替代，因此我们认为这一高分子量的凝集素样蛋白其活性可能不依赖于钙离子。

总之，我们从烙铁头蛇毒腺中克隆到的 4 个凝集素样蛋白基因，其编码的氨基酸序列中均存在 CRD 结构，与已报道的其他蛇毒凝集素样蛋白的氨基酸序列有很高的同源性。这项工作为研究蛇毒凝集素样蛋白的多样性和基因进化提供了有益的资料，对深入研究蛇毒凝集素样蛋白结构和功能的关系及蛇毒凝集素样蛋白基因的表达打下了良好的基础。

参考文献：

Atoda H, Hyuga M, Morita T. 1991. The primary structure of coagulation factor Ⅸ/factor Ⅹ-binding protein isolated from the venom of *Trimeresurus flavoviridis*: Homology with asialoglycoprotein receptors, proteoglycan core protein, tetranectin, and lymphocyte Fc epsilon receptor for immunoglobulin E[J]. *J. Biol. Chem.*, **266**: 14903-14911.

Atoda H, Ishikawa M, Mozuno H, *et al*. 1998. Coagulation factor Ⅹ-binding protein from *Deinagkistrodon acutus* venom is a Gla domain-binding protein[J]. *Biochemistry*, **37**: 17361-17370.

Chen Y L, Tsai K W, Chang T, *et al*. 2000. Glycoprotein Ib-binding

protein from the venom of *Deinagkistrodon acutus*-cDNA sequence, functional characterization, and three-dimensional modeling [J]. *Thromb. Haemost.*, **83**: 119 - 126.

Chung C H, Au L C, Huang T F. 1999. Molecular cloning and sequence analysis of aggretin, a collagen-like platelet aggregation inducer[J]. *Biochem. Biophys. Res. Commun.*, **263**: 723 - 727.

Francischetti I M B, Saliou B, Leduc M, *et al*. 1997. Convulxin, a potent platelet aggregating protein from *Crotalus durissus terrificus* venom, specially binds to platelets[J]. *Toxicon*, **35**: 1217 - 1228.

Fukuda K, Mizuno H, Atoda H, *et al*. 2000. Crystal structure of flavocetin-A, a platelet glycoprotein Ib-binding protein, reveals a novel cyclic tetramer of C-type lectin-like heterodimers[J]. *Biochemistry*, **39**: 1915 - 1923.

Fujimura Y, Titani K, Usami Y, *et al*. 1991. Isolation and chemical characterization of two structurally and functionally distinct forms of botrocetin, the platelet coagglutin isolated from the venom of *Bothrops jararaca*[J]. *Biochemistry*, **30**: 1957 - 1964.

Leduc M, Bon C. 1998. Cloning of subunits of convulxin, a collagen-like platelet-aggregating protein from *Crotalus durissus terrificus* venom[J]. *Biochem. J.*, **333**: 389 - 393.

Matsuzaki R, Yoshiara E, Yamada M, *et al*. 1996. cDNA cloning of Ⅸ/Ⅹ-bp, a heterogeneous two-chain anticoagulant protein from snake venom[J]. *Biochem. Biophys. Res. Commun.*, **220**: 382 - 387.

Mizuno H, Fujimoto Z, Koizumi M, *et al*. 1997. Structure of coagulation factors Ⅸ/Ⅹ-binding protein, a heterodimer of C-type lectin domains[J]. *Nature Struct. Biol.*, **4**: 438 - 441.

Mizuno H, Fujimoto Z, Koizumi M, *et al*. 1999. Crystal structure of coagulation factor Ⅸ-binding protein from habu snake venom at 2.6Å: Implication of central loop swapping based on deletion in the linker region[J]. *J. Mol. Biol.*, **289**: 103 - 112.

Peng M, Lu W, Beviglia L. 1993. Echicetin: A snake venom protein that inhibits binding of von Willebrand factor and alboaggregins to platelet glycoprotein Ib[J]. *Blood*, **81**: 2321-2328.

Sakurai Y, Fujimura Y, Kokubo T, *et al*. 1998. The cDNA cloning and molecular characterization of a snake venom platelet glycoprotein Ib-binding protein, mamushigin, from *Agkistrodon halys blomhoffii* venom[J]. *Thromb. Haemost.*, **79**: 1199 - 1207.

Sen U, Vasudevan S, Subbarao G, *et al*. 2001. Crystal structure of the von Willebrand factor modulator botrocetin[J]. *Biochemistry*, **40**: 345 - 352.

Shin Y, Morita T. 1998. Rhodocytin, a functional novel platelet agonist belonging to the heterodimeric C-type lectin family, induces platelet aggregation independently of glycoprotein Ib[J]. *Biochem. Biophys. Res. Commun.*, **245**: 741 - 745.

Shin Y, Okuyama I, Hasegawa J, *et al*. 2000. Molecular cloning of glycoprotein Ib-binding protein, flavocetin-A, which inhibits platelet aggregation[J]. *Thromb. Res.*, **99**: 239 - 247.

Taniuchi Y, Kawasaki T, Fujimura Y, *et al*. 1995. Flavocetin-A and -B, two high molecular mass glycoprotein Ib-binding proteins with high affinity purified from *Trimeresurus flavoviridis* venom, inhibit platelet aggregation at high shear stress[J]. *Biochim. Biophys. Acta*, **1244**: 331 - 338.

Usami Y, Fujimura Y, Suzuki M, *et al*. 1993. Primary structure of two-chain botrocetin, a von Willebrand factor modulator purified from the venom of *Bothrops jararaca* [J]. *Proc. Natl. Acad. Sci. USA*, **90**: 928 - 932.

Usami Y, Suzuki M, Yoshida E, *et al*. 1996. Primary structure of alboaggregin-B purified from the venom of *Trimeresurus albolabris*[J]. *Biochem. Biophys. Res. Commun.*, **219**: 727 - 733.

Zingali R B, Jandrot-Perrus M, Guillin M C, *et al*. 1993. Bothrojaracin, a new thrombin inhibitor isolated from *Bothrops jararaca* venom: Characterization and mechanism of thrombin inhibition[J]. *Biochemistry*, **32**: 10794 - 10802.

简 讯

山东省发现文须雀

贾少波[1]

（南京大学 生命科学学院，江苏 南京 210093；聊城大学 生物学系，山东 聊城 252059）

笔者分别于1994、1995、1998和2000年在山东省聊城市观察到文须雀（*Panurus biarmicus*）种群。观察到的具体时间、地点和数量见表1。

表1 在山东省观察到的文须雀的情况

时间	地点	数量（只）
1994-11-11	聊城市东昌湖岸芦苇灌草丛	♂22 ♀20
1995-10-28	聊城市东昌湖岸芦苇灌草丛	♂9 ♀12
1998-03-01	阳谷县寿张镇黄河金堤北面脚下芦苇灌草丛	♂14 ♀8
2000-02-25	聊城市东昌湖岸芦苇灌草丛	♂5 ♀4

该种鸟类活动特点是：成小群（几只至40几只）从高空迅速飞下，落于湖边或沼泽地，在水边芦苇和灌草丛中取食。活动范围自地面至芦苇上端，尤其喜在中下部活动。性活泼，伴有小声的喧闹鸣叫。往往集中在一个10~20 m^2的小区域，活动20~30 min再换一个类似区域。文献记录该种在我国分布于河北以北地区，包括新疆、青海、甘肃、内蒙和东北。该种鸟为山东分布的新纪录。

1. 通讯联系人，Tel：025-3595389，E-mail：sbjia@sina.com